SYMPOSIUM

Analyzing Big Data: Social Choice and Measurement

John W. Patty, *Washington University in St. Louis*
Elizabeth Maggie Penn, *Washington University in St. Louis*

Social scientists commonly construct indices to reduce and summarize complex, multidimensional data. Examples include assigning texts to topical categories, converting roll call matrices into "ideal points," producing brain images from raw fMRI data (Aguirre 2014), various node- and edge-indices in network analysis, and the wide array of composite indicators that seek to distill potentially thorny concepts such as income inequality, democratization, civic competence, and environmental responsibility into a list of numbers. The benefits of these indices are clear: a single number is simple to interpret, and such indices can facilitate communication among scholars, policymakers, and the public on important issues (see Nardo et al. 2005). Moreover, indices make it possible to track the absolute and relative performance of objects over time, and lend themselves to empirical study as both dependent and independent variables because they can be readily incorporated into a regression.

In this contribution we begin by noting that empirical analysis of many interesting social science topics—and in particular those popularly associated with "big data" such as networks, text analysis, and genetics—always require data reduction. Data reduction is achieved through aggregating higher-dimensional data into lower-dimensional measures. In other words, we conceive of the "bigness" of data as a function of the data's underlying conceptual structure. Much more than being a function of the number of observations or the number of variables, data is "big" if the concepts underlying the data—the data's *raison(s) d'être*—are more complicated than a list of vectors of numbers. Social networks, texts, genomes, and brain scans all satisfy this requirement.

With this notion of big data in hand, our argument is that formal theory is central to modern data analysis. Furthermore, our argument is that formal theory is the heart of measurement: regardless of one's ultimate approach to "theory meets data," à la "Empirical Implications of Theoretical Models," big data concepts are generally unamenable to classical, unidimensional empirical analysis, and this reality strikes prior to even reaching traditional questions of causal identification. In other words, while one might think of formal theory as describing an apparatus that might produce hypotheses to be "tested" by the data, we point out here that formal theory, specifically *social choice theory*, speaks to how the data with which the test will be carried out was created. Social choice theory is central to measurement because it is primarily concerned with aggregation: the creation of a measure of some underlying concept (e.g., social welfare, majority will, inequality, power) from a set of potentially heterogeneous inputs (e.g., individual preferences, ballots, incomes, relative capabilities).[1]

As the availability of detailed and complex data—and the computational resources used to analyze it—has surged in recent years, the need to reduce the dimensionality of complex objects for theoretical and practical study has grown. Consider, for example, the class of social networks, which are highly multidimensional and complex objects. Directly analyzing networks per se would require analyzing each potential network structure separately; networks are not ordered objects, so there are no simple "greater than / less than" relationships that could be assumed. The size of the space of possible networks makes the analysis of all possible network structures impractical for all but very small networks.[2]

Putting these impracticalities aside, there is a more important reason for why one would not want to consider all possible network structures. Such an analysis might tell us that certain networks matter and others do not, but it would provide no insight into how or why networks matter. Answering these questions requires a theory of what network characteristics matter. From the "how" perspective, a theory is required to render the analysis portable, or applicable to data not in the sample. From the "why" perspective, a theory is required to make the analysis extendable to topics outside the domain of the immediate question at hand. Because of the complexity of the space of possible networks, any meaningful theory of networks must involve data reduction through *aggregation*. In a social network, for example, aggregation could involve transforming information about every individual's connection to every other individual into a ranking of each person's "connectedness" within the network. For any type of transformation of this sort, there *must* exist instances in which two different networks will yield the same list of "connectedness" rankings, because there are more potential networks than there are rankings. In these situations the connectedness ranking deems the two different networks equivalent. Phrased differently, if a set of networks generate identical connectedness rankings, then differences between those networks are not *measurable* with respect to the ranking algorithm being utilized.

Any process of data reduction necessarily involves choices about measurement: the information to retain versus the information to lose. As mentioned previously, many emerging and classic areas of study rely on aggregation in their empirical

doi:10.1017/S1049096514001814

analysis. While we use the running example of networks in this article, our point is applicable to any problem of reducing multidimensional data into a unidimensional measure—and as we have argued previously, these problems are unavoidable in "big data" analyses.

NETWORK ANALYSIS AND AGGREGATION

To illustrate some of our arguments we focus on the topic of network analysis, which has played a growing role across the various subfields of political science[3] in recent years with examples being coordination and collective action,[4] legislative behavior,[5] judicial decision-making,[6] policy adoption,[7] causal inference,[8] and the formation of political opinions and attitudes.[9] Networks are big data in the truest sense: they are not expressible in a unidimensional fashion. Accordingly, analysis (e.g., regression analysis or simple reporting) usually requires some data reduction as an intermediate step (e.g., Frankel and Reid 2008).[10] Because of this practical challenge, social choice/axiomatic approaches to theory development in political science are relevant for this new area of research. Put succinctly, network analysis—like all big data enterprises—is necessarily reliant on (typically multiple) theoretical choices with respect to measurement; social choice theory is primarily concerned with the logical implications of such choices.

A Running Example: Centrality Indices

A *network* is pair of nodes (e.g., people), N, and edges (e.g., connections), $E \subset N \times N$. We denote a network by G = (N, E). A *centrality index* is a function, c, that assigns every node in a network a score that represents how central the node is relative to all of the other nodes in the network.[11] There are many such indices, but we focus on three indices that are widely used, which we define informally:[12]

1. The *degree centrality* of a node, denoted by c_d, is the number of other nodes to which the node in question is connected.
2. The *betweenness centrality* of a node, denoted by c_b, is the number of shortest paths between *other* pairs of nodes that contain the node in question.
3. The *closeness centrality* of a node, denoted by c_c, is the average length of the shortest paths from the node to every other node.

While a centrality index assigns a real number to each node (i.e, it is a cardinal measure), we can also think about the index as returning a ranking of the nodes (i.e., as an ordinal construct).

AN EXAMPLE: THE FLORENTINE MARRIAGE NETWORK

The Florentine Marriage Network (Padgett and Ansell 1993) describes the marital ties among 15 major families in Florence in the early fifteenthth century. It is displayed in figure 1, and the rankings of the families for the three indices previously defined are displayed in table 1. As indicated by each of the three indices, the Medici family is clearly "central" in this network. However, the indices differ with respect to the rankings of the remaining families. Thus, the choice of which index to use in an analysis is potentially important: each of these indices aggregates the information available in the Florentine network into a smaller format (a list of 15 numbers). In so doing, of course, a centrality index loses some of the information contained in the original network. Each of these indices makes different choices about what information to lose.

To make this point tangible, suppose that one uses degree centrality, c_d, to represent the centralities of the families. In so doing, the analyst has equated the Florentine network as displayed in figure 1 with each of the three networks displayed in figure 2. Each of the three "degree-equivalent" networks assigns each of the families the same degree centrality as they possess in the network displayed in figure 1 (so that, for example, the Medici have the highest degree centrality in each of these three networks) but, as described in figure 2, each of these degree-equivalent networks produce a different family as the most "close" and most "between." That is, in each of the three networks, a family other than the Medici is assigned the highest c_c *and* c_b scores. However, remember that, with respect to the original (i.e., true) network structure, the Medici family should be assigned highest closeness centrality and the highest betweenness centrality. This divergence is because, by the definition of the data reduction/aggregation process used here (i.e., degree centrality), the three networks in figure 2 are not "measurable" (or distinguishable) with respect to degree centrality, but *are*

Figure 1
The Florentine Marriage Network

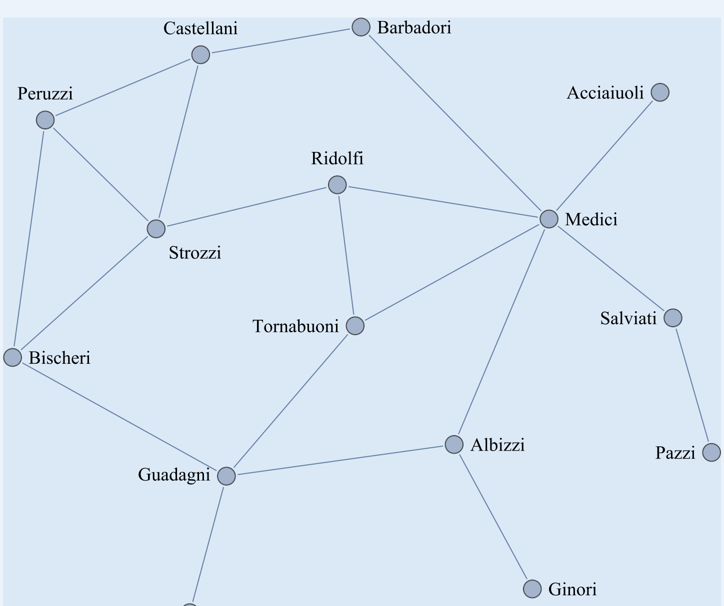


Table 1

Various Centrality Rankings of the Florentine Families in Figure 1

	Degree (cd)			Betweenness (cb)			Closeness (cc)	
1	Medici	6	1	Medici	47.5	1	Medici	0.56
2t	Guadagni	4	2	Guadagni	23.17	2	Ridolfi	0.5
2t	Strozzi	4	3	Albizzi	19.33	3t	Albizzi	0.48
4t	Bischeri	3	4	Salviati	13	3t	Tornabuoni	0.48
4t	Albizzi	3	5	Ridolfi	10.33	5	Guadagni	0.47
4t	Tornabuoni	3	6	Bischeri	9.5	6t	Barbadori	0.44
4t	Ridolfi	3	7	Strozzi	9.33	6t	Strozzi	0.44
4t	Peruzzi	3	8	Barbadori	8.5	8	Bischeri	0.4
4t	Castellani	3	9	Tornabuoni	8.33	9t	Salviati	0.39
10t	Salviati	2	10	Castellani	5	9t	Castellani	0.39

Note: "Degree," "Betweenness," and "Closeness" Refer to degree centrality, betweenness centrality, and closeness centrality, as defined previously in text.

measurable/distinguishable with respect to both closeness and betweenness centrality.

Figure 2 indicates how different indices "lose" or "forget" different aspects of the network: degree centrality does not differentiate between the original Florentine network and the networks in figure 2, whereas closeness and betweenness each do differentiate between them. Figure 3 contrasts the original Florentine network with a richer network—one containing a strictly larger set of links—that maintains the betweenness ranking of the 15 families. In the new network, the Guadagni family now is the "degree center," rather than the Medici family. This enlarged network demonstrates the information lost by the betweenness centrality index: it can not distinguish between the enlarged network (represented by the additional dashed lines) and the original one, unlike degree centrality.

Any procedure that takes "big" data and makes it "smaller" is aggregating (and hence potentially losing) information and, in the process, making choices about what to be responsive to versus what to ignore.

Spanning fields such as sociology, economics, physics, and, of course, political science, many scholars have noted that different centrality indices measure different things; space precludes us from a rote and necessarily incomplete recounting of the many such published notices of this point. However, less well appreciated is that this same point has been made regarding social choice. Arrow's Impossibility Theorem (Arrow 1963) implies that different aggregation methods of individual preferences (i.e., different voting systems) "measure different things." More specifically, in a general and precise sense, there is no perfect aggregation method for preferences— we are simply and humbly linking this insight with the received wisdom regarding centrality indices in the study of networks.

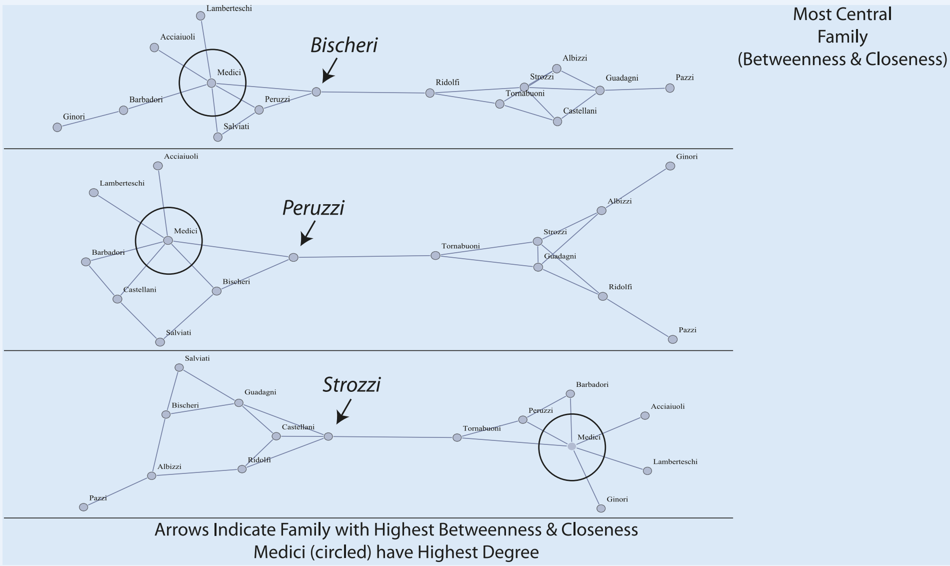
The link between Arrow's Impossibility Theorem and the notion that different centrality indices (or, more generally, data reduction techniques) measure different theoretical notions *is* the foundation of the equivalence of the "big data measurement problem" with the central question of social choice theory: *how can we measure social welfare or preference*? The choices that an analyst must make and the trade-offs inherent in choosing indices to represent a network or other "big data" are the same as those faced by a social planner attempting to aggregate individual preferences, economic factors, and other criteria to produce a social preference over a set of alternative policies. As we have argued at length elsewhere, "social choice theory informs us about the possibilities and impossibilities of aggregation. Furthermore, and tellingly, aggregation is simply that: putting various things together to produce an output" (Patty and Penn 2014b, 7). Any procedure that takes "big" data and makes it "smaller" is aggregating (and hence potentially losing) information and, in the process, making choices about what to be responsive to versus what to ignore.

AXIOMATIC CHARACTERIZATIONS OF CENTRALITY

Having worked through an available and palpable example of data reduction in a big data problem, we examine a few examples of how social choice results and axiomatic analysis can illustrate these measurement problems. Precisely because aggregation is valuable when one is confronting complicated, high-dimensional data, practically gauging the performance and characteristics of an aggregation approach through trial and error is cumbersome. An axiomatic approach—deriving a set of desiderata one would like the aggregation method to satisfy and then characterizing the (possibly empty) class

Figure 2

Three Degree-Equivalents of the Florentine Network



of methods that satisfy them—is valuable in these situations because of its generality and because it directly and clearly confronts the questions of measurability that we discussed. In this section, we demonstrate the value of this approach while using our running example of centrality indices.[13] We start by defining two axioms that one might want a centrality index to satisfy; in the process we discuss the close connection between these axioms (defined for network data) and well-known social choice theoretic axioms that characterize properties of voting systems (defined for ballot data).[14]

Axiom 1: Positive Responsiveness

A centrality index satisfies *positive responsiveness* if it maintains or increases the ranking of nodes that have gained more links relative to nodes that have not. To illustrate more formally, suppose that node i was ranked (weakly) higher than node j for a network G. Now add one extra edge between i and an $h \neq j$ that did not exist before, and call this new network G'. If the index is positively responsive then it must now rank i as *strictly* better than j. Positive responsiveness is a well-known axiom applied to voting systems. It says that if the voting rule originally ranked x as weakly better than y and a person changes their ballot to strictly improve x relative to y then then rule must rank x *strictly* better than y after this change occurs.

Positive responsiveness is directly related to measurement in the sense that an index that satisfies it always responds in a sensible monotonic fashion to a simple and easily discernible alteration of the network. In particular, focusing on the notion of centrality, positive responsiveness requires that the index never deem a node less central as the result of that node adding links to more nodes. Insofar as—for theoretical reasons—one thinks a node "being connected to" more nodes never makes the node in question less central to the network, then the appropriate measure of centrality *must* satisfy positive responsiveness.

In social choice and mechanism design, positive responsiveness is a type of *monotonicity* axiom, which describes a system in which "moving a candidate up in one's ballot" never makes the candidate less likely to win. While there are generalizations of this idea, it is well-known that some version of monotonicity is required to ensure that no individual has an incentive to "vote strategically" (see Muller and Satterthwaite 1997). This link with positive responsiveness provides yet another indication of the connection between the axiomatic approach and measurement: if an index fails to satisfy positive responsiveness, then there will be some situations in which "a node seeking to maximize its centrality" would have an incentive to "disconnect itself" from one or more other nodes.

Axiom 2: Independence of Irrelevant Edges

An index satisfies *independence of irrelevant edges* (IIE) if it produces the same relative ranking of nodes i and j for any two graphs in which the edges containing i and j have not changed. For example, suppose that G and G' are two networks where i has the same set of edges in both G and G' and j has the same set of edges in both G and G'. Then if i is ranked above j in G it should also be ranked above j in G', and adding an edge between two different nodes, a and b, should not affect the relative i, j ranking. Readers familiar with Kenneth Arrow's famous impossibility theorem—the most well-known result in social choice theory—will note the similarity between this condition and Arrow's *independence of irrelevant alternatives*, which states that a voting system's relative ranking of alternatives x and y should be invariant to how some other alternative z fares in each voter's ballot.

The IIE axiom is less intuitive than positive responsiveness, but it too is directly linked to measurement. Whereas positive responsiveness specifies what an index *should* measure, IIE specifies what an index *should not* measure. An index that satisfies IIE can rank a node only on the basis of the nodes to which that node is attached. Accordingly, the IIE axiom is directly about data reduction: any index that satisfies it can be accurately calculated for subnetworks, as displayed in figure 4. The figure displays a situation in which only a subset of the nodes of the original Florentine network are scored. If the links denoted by solid lines are used to rank the nodes labeled in boldface, the relative rankings of those nodes will remain unchanged if the index satisfies IIE. Put another way, an index that violates IIE in some cases will yield faulty rankings of perfectly observed nodes as a result of faulty data regarding other nodes. In big data settings, such as social networks and text analysis, the premise that one has collected perfect, or "equal-quality," data about every unit of analysis is generally implausible. Using an index that satisfies IIE mitigates some of an analyst's worries about using such data.

Theorem 1 (*C.f.*, van den Brink and Gilles 2003)

A centrality index satisfies Positive Responsiveness and IIE if and only if it is equivalent to degree centrality, cd. That is, for any pair of nodes i *and* j, i *is ranked higher by the index than* j *if and only if* i *has strictly more links than* j *does.*

Theorem 1 states that requiring a centrality index to satisfy positive responsiveness and IIE implies that one must use an index that is equivalent to degree centrality, c_d.[16] Accordingly, the examples and discussion regarding the Florentine Marriage Network illustrate that closeness centrality and betweenness centrality *must* each violate at least one of these axioms.[17]

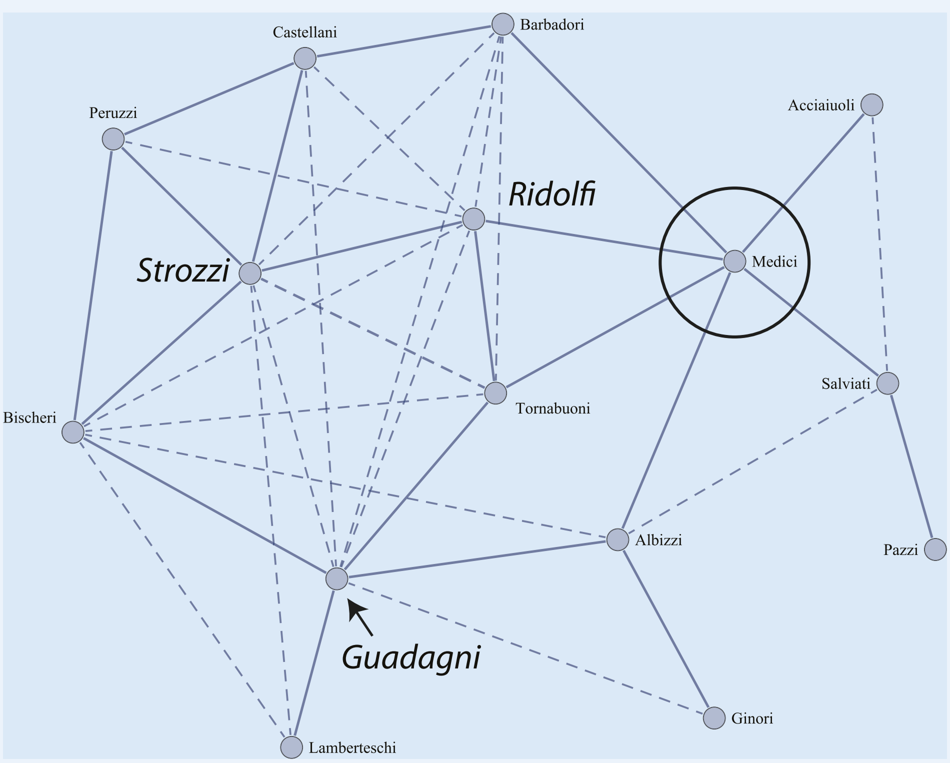
In big data settings, such as social networks and text analysis, the premise that one has collected perfect, or even "equal-quality," data about every unit of analysis is generally implausible. Using an index that satisfies IIE mitigates some of an analyst's worries about using such data.

We are now in a position to state an interesting characterization theorem, which is a straightforward extension of a theorem recently proved by van den Brink and Gilles (2003).[15]

This point bears repeating: the *practical* import of Theorem 1 is that it highlights for analysts the conceptual and theoretical properties that they want to "measure," and—in the best of cases—identifies the class of data reduction techniques (in this case, centrality indices) that achieves the analysts' goals. Put this way, Theorem 1 clearly lays out the following advice to a practical analyst: *if the theoretical concept you are trying to measure respects positive responsiveness and IIE, then it is precisely and uniquely measured by degree centrality*. Note that this conclusion is powerful and—more importantly—the power of the conclusion derives not from some sort of high-powered math under the proof of Theorem 1 but, rather, from the clarity and precision emanating from the abstract nature of the IIE and positive responsiveness axioms. Simply put, the fullest power of

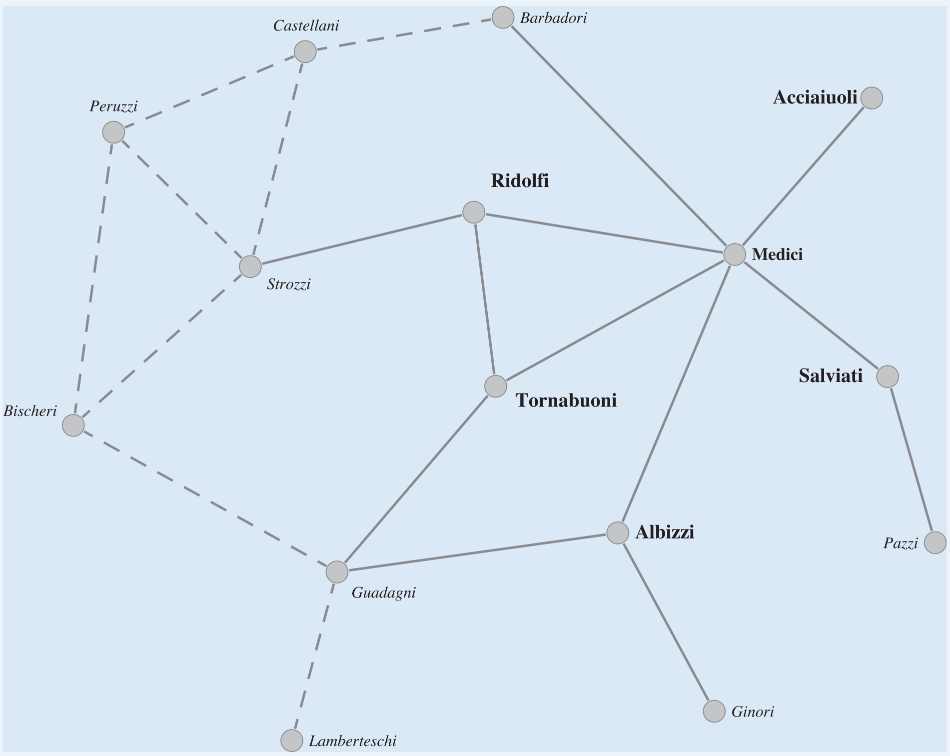
Figure 3

An Enlarged, Betweeness-Equivalent of the Florentine Network



Enlarged Florentine Network, New Links are Dashed Arrow Indicates Family with Highest Degree Italicized Families have Higher Degrees than Medici (circled).

Figure 4
Subnetworks and IIE



Sub-Network of Florentine Network, "Omitted" Links are Dashed Only Bold-Faced Families Scored, Italicized Families Not Scored
IIE Implies Relative Rankings of Bold-Faced Families Unchanged.

formal theory is derived from the clarity of its starting points. Axioms, definitions, assumptions: when applied correctly, these foundations of formal theory are identical to the gold standards of transparency and replicability in empirical work. This is no coincidence: "data" is nothing until we use it to measure something, and "measurement" is accordingly *inherently* theoretical. Our argument in this contribution is that, while nothing is perfect, axiomatic approaches to measurement are the most reliable theoretical approaches available.

CONCLUSION

In this article, we have attempted to clearly illustrate the fundamental connections between social choice and *measurement* in social science. Especially with the dawning of the era of big data, data reduction—and hence aggregation—is both practically necessary and done with increasing frequency. While we stand on the shoulders of giants—many, many scholars in essentially every social science have noted the importance of theoretically grounded measurement—the rapidly growing complexity of the data being explored and theories being explored increase the value of the axiomatic approach to measurement that social choice has been founded on since the seminal contributions of Arrow, Gibbard, and Satterthwaite (Arrow 1963; Gibbard 1973; Satterthwaite 1975). Divining the properties of any specific aggregator, such as a centrality index through trial and error, Monte Carlo methods, or other search methods, is time consuming and necessarily incomplete—and even more so as the curse of dimensionality comes into even starker relief with the emergence of big data. While axiomatic methods are by no means a panacea—axioms are useful precisely because they are stark, and characterization/impossibility results offer maximum validity at the price of nuance[18]—they represent a clear complement to even the best approaches generally employed by empirical researchers confronting complex data.

ACKNOWLEDGMENTS

We thank Bill Clark, Matt Golder, Jacob Montgomery, and Betsy Sinclair for incredibly useful and generous feedback while writing this. All remaining errors are, of course, our own. ■

NOTES

1. In Patty and Penn (2014b), we use this point as the basis for linking the classic "impossibility theorems" of social choice theory (e.g., Arrow 1963; Gibbard 1973; Satterthwaite 1975) with the role of explanations and justifications in constraining and legitimizing democratic governance.
2. For example, consider the space required for a recent strategic analysis of network structure and information in Patty and Penn (2014a), where we consider only the set of 3-node networks.
3. See McClurg and Young (2011); Ward, Stovel, and Sacks (2011); and Lazer ([2011) for recent summaries, as well as Heaney and McClurg – ([2009), Huckfeldt (2009), Siegel (2011) and Hafner-Burton, Kahler, and Montgomery (2009) for more subfield-focused reviews.
4. Among others, see Mutz (2002); and McCubbins, Paturi, and Weller (2009).
5. For example, Fowler (2006b,a), Zhang et al. (2008), and Cho and Fowler (2010).
6. One example among many is Fowler, Johnson, Spriggs II, Jeon, and Wahlbeck (2007).
7. A small selection from a large literature includes Rhodes (1990), Mintrom and Vergari (1998), and Simmons, Dobbin, and Garrett (2007).
8. See Fowler, Heaney, Nickerson, Padgett, and Sinclair (2011).
9. Exemplary of a large literature are McClurg (2006) and Klofstad, Sokhey, and McClurg (2013).
10. For example, even drawing a network requires choices to be made: a nearly ubiquitous example of this is the Fruchterman-Reingold algorithm for drawing networks as 2-dimensional graphs (Fruchterman and Reingold 1991).
11. Similar indices can be defined for edges in the network as well.

12. Space precludes a lengthy treatment of such indices; the interested reader is referred to Borgatti (2005), Borgatti and Everett (2006), Jackson (2008), Newman (2010), and Ward, Stovel, and Sacks. (2011).
13. Note that van den Brink and Gilles' axioms and corresponding theorem that follow (van den Brink and Gilles 2003) are defined for the more general case of directed graphs. Their proof technique holds for the case of undirected graphs, and to simplify our discussion we have modified their axioms and results slightly to accommodate this less general class of graphs that we consider here.
14. For reasons of space, we omit a third axiom that is satisfied by all centrality indices that are used. This axiom, anonymity, requires that the labels of the nodes have no effect on the ranking produced by the index. Anonymity of a network centrality index is closely related to both the anonymity and neutrality axioms for voting systems, which require the system to give every voter equal weight in determining the outcome and, respectively, to treat each candidate equally.
15. For the reasons discussed in note 14 we are imposing anonymity in the definition of a centrality index.
16. The interested reader will note the similarity between this result and May's Theorem (May 1952), a famous social choice-theoretic result that proves that the only anonymous, neutral and positively responsive voting rule is plurality rule.
17. Closeness centrality and betweenness centrality are each anonymous.
18. For example, Theorem 1 offers no guidance about questions such as "how dependent on irrelevant edges is closeness centrality?"

REFERENCES

Aguirre, Geoffrey K. 2014. "Functional Neuroimaging: Technical, Logical, and Social Perspectives." *Hastings Center Report* 44 (s2): S8–S18.

Arrow, Kenneth J. 1963. *Social Choice and Individual Values, 2nd ed.* New York: Wiley.

Borgatti, Stephen P. 2006. "Centrality and Network Flow." *Social Networks* 27 (1): 55–71.

Borgatti, Stephen P., and Martin G Everett. 2006. "A Graph-Theoretic Perspective on Centrality." *Social Networks* 28 (4): 466–84.

Cho, Wendy K. Tam, and James H. Fowler. 2010. "Legislative Success in a Small World: Social Network Analysis and the Dynamics of Congressional Legislation." *The Journal of Politics* 72 (1): 124–35.

Fowler, James H. 2006a. Connecting the Congress: A Study of Cosponsorship Networks. *Political Analysis* 14 (4): 456–87.

———. 2006b. "Legislative Cosponsorship Networks in the US House and Senate. *Social Networks* 28 (4): 454–65.

Fowler, James H. and Timothy R. Johnson, James F. Spriggs II, Sangick Jeon, and Paul J. Wahlbeck. 2007. "Network Analysis and the Law: Measuring the Legal Importance of Supreme Court Precedents." *Political Analysis* 15 (3): 324–46.

Fowler, James H., Michael T. Heaney, David W. Nickerson, John F. Padgett, and Betsy Sinclair. 2011. "Causality in Political Networks." *American Politics Research* 39 (2): 437–80.

Frankel, Felice, and Rosalind Reid. 2008. "Big Data: Distilling Meaning from Data." *Nature* 455 (7209): 30.

Fruchterman, Thomas M. J., and Edward M Reingold. 1991. Graph Drawing by Force-Directed Placement. *Software: Practice and Experience* 21 (11): 1129–64.

Allan Gibbard, Allan. 1973. "Manipulation of Voting Schemes: A General Result." *Econometrica* 41 (4): 587–601.

Hafner-Burton, Emilie M., Miles Kahler, and Alexander H. Montgomery. 2009. "Network Analysis for International Relations." *International Organization* 63 (03): 559–92.

Heaney, Michael T., and Scott D. McClurg. 2009. "Social Networks and American Politics: Introduction to the Special Issue." *American Politics Research* 37 (5): 727–41.

Huckfeldt, Robert. 2009. "Interdependence, Density Dependence, and Networks in Politics." *American Politics Research* 37 (5): 921–50.

Jackson, Matthew O. 2008. *Social and Economic Networks.* Princeton, NJ: Princeton University Press.

Klofstad, Casey A., Anand Edward Sokhey, and Scott D McClurg. 2013. "Disagreeing about Disagreement: How Conflict in Social Networks Affects Political Behavior." *American Journal of Political Science* 57 (1): 120–34.

Lazer, David. 2011. "Networks in Political Science: Back to the Future." *PS: Political Science and Politics* 44 (1): 61–68.

May, K. O. 1952. "A Set of Independent Necessary and Sufficient Conditions for Simple Majority Decision." *Econometrica* 20 (4): 680–84.

McClurg, Scott D. 2006. "The Electoral Relevance of Political Talk: Examining Disagreement and Expertise Effects in Social Networks on Political Participation." *American Journal of Political Science* 50 (3): 737–54.

McClurg, Scott D. and Joseph K Young. 2011. "Political Networks." *PS: Political Science and Politics* 44 (1): 39–43.

McCubbins, Mathew D., Ramamohan Paturi, and Nicholas Weller. 2009. "Connected Coordination: Network Structure and Group Coordination." *American Politics Research* 37 (5): 899–920.

Mintrom, Michael, and Sandra Vergari. 1998. " Networks and Innovation Diffusion: The Case of State Education Reforms." *Journal of Politics* 60: 126–48.

Muller, Eitan, and Mark A. Satterthwaite. 1977. "The Equivalence of Strong Positive Association and Strategy-Proofness." *Journal of Economic Theory* 14 (2): 412–18.

Mutz, Diana C. 2002. "The Consequences of Cross-Cutting Networks for Political Participation." *American Journal of Political Science* 46 (4): 838–55.

Nardo, Michela Michaela Saisana, Andrea Saltelli, Stefano Tarantola, Anders Hoffman, and Enrico Giovannini. 2005. *Handbook on Constructing Composite Indicators: Methodology and User Guide.* Technical report, OECD Publishing, Paris, France.

Newman, Mark. 2010. *Networks: An Introduction.* New York: Oxford University Press.

Padgett, John F., and Christopher K Ansell. 1993. "Robust Action and the Rise of the Medici, 1400–1434." *American Journal of Sociology* 98 (6): 1259–1319.

Patty, John W. and Elizabeth Maggie Penn. 2014a. "Sequential Decision-Making and Information Aggregation in Small Networks." *Political Science Research & Methods* 2 (2): 243–71.

———. 2014b. *Social Choice and Legitimacy: The Possibilities of Impossibility.* New York: Cambridge Universsity Press.

Rhodes, R. A. W. 1990. "Policy Networks." *Journal of Theoretical Politics* 2 (3): 293.

Satterthwaite, Mark A. 1975. "Strategy-Proofness and Arrow's Conditions: Existence and Correspondence Theorems for Voting Procedures and Social Welfare Functions." *Journal of Economic Theory* 10 (2): 187–217.

Siegel, David A. 2011. "Social Networks in Comparative Perspective." *PS: Political Science and Politics* 44 (1): 51–54.

Simmons, Beth, Frank Dobbin, and Geoffrey Garrett. 2007. "The Global Diffusion of Public Policies: Social Construction, Coercion, Competition, or Learning.? *Annual Review of Sociology* 33 (1): 449–472.

van den Brink, Rene, and Robert P. Gilles. 2003. "Ranking by Outdegree for Directed Graphs." *Discrete Mathematics* 271(1): 261–70.

Ward, Michael D., Katherine Stovel, and Audrey Sacks. 2011. Network Analysis and Political Science. *Annual Review of Political Science* 14: 245–64.

Zhang, Y., and A. J. Friend, A. L. Traud, M. A. Porter, J. H. Fowler, and P. J. Mucha. 2008. "Community Structure in Congressional Cosponsorship Networks." *Physica A: Statistical Mechanics and its Applications* 387 (7): 1705–12.